프로젝션 함수와 허프 변환을 이용한 눈동자 중심점 찾기*

최연석* · 문원호* · 김철기** · 차의영*

*부산대학교 컴퓨터공학과

Detection of Pupil Center using Projection Function and Hough Transform

Yeon-seok Choi* · Won-ho Mun* · Cheol-ki Kim** · Eui-young Cha*

*Dept. of Computer Engineering, Pusan National University

**Dept. of Design, Pusan National University

E-mail : yeonseokchoi@pusan.ac.kr

요 약

본 논문에서는 프로젝션 함수와 허프 변환을 이용하여 영상에서 눈동자를 찾는 방법을 제안한다. 먼저, 영상으로부터 얼굴영역을 추출한 다음, 눈썹과 눈동자의 밝기변화의 특징을 이용할 수 있는 integral projection function과 variance projection function을 사용하여 눈 영역을 검출한다. 검출된 눈 영역에서 눈동자 중심좌표를 구하기 위해 원형 허프 변환을 이용한다. 원형 허프 변환에 사용된 좌표는 sobel edge mask를 사용하여 구한다. FERET database의 정면 얼굴 영상을 이용하여 제안된 방법으로 실험한 결과 만족할 만한 결과가 나왔다.

ABSTRACT

In this paper, we proposed a novel algorithm to detect the center of pupil in frontal view face. This algorithm, at first, extract an eye region from the face image using integral projection function and variance projection function. In an eye region, detect the center of pupil positions using circular hough transform with sobel edge mask. The experimental results show good performance in detecting pupil center from FERET face image.

키워드

eye tracking, eye detection, pupil detection, Hough transform

Ⅰ. 서 론

9.11 테러사건 이후 테러에 대비한 보안문제가 최근 대두되고 있다. 테러사건 뿐만 아니라 생체정보를 필요로 하는 보안 기술이 주목 받고 있다. 생체정보는 타인에 의해 쉽게 복제 될 수 없는 개인의 고유한 특징이며, 이러한 정보를 이용해서 개인을 인식하는 기술에는 지문, 서명검증, 음성, 얼굴 인식 등이 있다[1]. 특히 카메라를 통해 영상에서 사람의 얼굴을 검출하는 것은 로봇, 군사, 감시 분야에서 널리 응용되고 있으며 최근 사람 행동 패턴 분석 등 과 같이 사람과 함께 하는 로봇이 필요로 함에 따라 얼굴 인식, 표정 인식 등과 같은 연구들이 활발히 진행되고 있다[2].

얼굴 인식, 표정 인식 등과 같은 얼굴 영상을 이용한 연구들은 일반적으로 입력 영상에 대한 정규화 과정이 필요하게 되는데 정규화는 주로 눈의 위치를 기준으로 이루어진다. 그렇기 때문에 눈의 위치에 대한 정보는 필수적인 정보이다. 얼굴 영상에서 눈의 위치를 검출하기 위하여 많은 연구들이 진행되어 오고 있다. A.L. Yuille 등은 템플릿 정합 방법을 이용하여 눈 검출을 시도하였는데[3], 이 방법은 수행시간이 오래 걸리는 단점이 있다. A. Pentland 등은 Eigenspace 방법을 이용하여 눈 검출을 시도하였다[4]. 그러나 이 방법은 학습 데이터 영상에 큰 영향을 받는 단점이

* 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임. (No. 2010-00282242)

있다. G.C. Feng 등은 variance projection function(VPF)를 이용하여 눈 검출을 시도하였다[5]. 이 방법은 눈이 작거나 감고 있는 영상에서는 검출하기 어려운 단점이 있다. 기존의 방법들은 자세, 표정, 조명 등의 변화에 잘 대처하지 못하거나 학습 데이터 영상에 민감한 문제점 들이 있다.

따라서 본 논문에서는 먼저 AdaBoost를 이용하여 얼굴영역을 검출한 다음[6], G.C. Feng 등이 제시한 VPF를 이용하여 눈의 후보영역과 눈동자 중심점을 찾는다. 검출된 눈동자 좌표는 눈동자의 특징을 이용하여 검증한다. 만약, 검출된 눈동자 명암도 값이 평균 피부 명암도보다 작으면 검증을 통과하며, 그렇지 않은 경우 후보영역에서 원형 허프 변환(Circular Hough transform)을 이용하여 눈동자 중심을 찾는다.

II. 얼굴영역 검출 및 1차 눈동자 검출

눈동자 검출을 하기에 앞서, 입력된 영상에서 얼굴영역을 먼저 검출해야 한다. 입력된 영상에서 얼굴 영역을 검출하기 위하여 본 논문에서는 AdaBoost 학습 알고리즘을 이용하며[6], 분류기는 Haar-like feature를 사용하였다[7].

두 번째로, 검출한 얼굴 영역에서 눈 영역을 포함한 후보 영역을 검출한다. 후보 영역 검출하기에 앞서 사람의 얼굴의 특징을 이용하여 왼쪽 눈 영역과 오른쪽 눈 영역의 범위를 축소할 수 있다. 본 논문에서는 얼굴 영역 부분을 4등분하여 왼쪽 위 부분을 오른쪽 눈 영역, 오른쪽 부분을 왼쪽 눈 영역으로 설정하였다.

축소된 눈 영역에서 먼저 integral projection function(IPF)를 구한다. 식(1), 식(2)에서 $I(x,y)$는 (x,y) 위치에서 명암 값이고, $V(x)$는 영역 $[y_1,y_2]$의 명암도의 합이며, $H(y)$는 영역$[x_1,x_2]$의 명암도의 합이다.

$$V(x)=\int_{y_1}^{y_2} I(x,y)dy \tag{1}$$

$$H(y)=\int_{x_1}^{x_2} I(x,y)dx \tag{2}$$

$V(x)$ 와 $H(y)$의 평균 IPF는 식(3), 식(4)와 같다.

$$V_m(x)=\frac{1}{y_2-y_1}\int_{y_1}^{y_2} I(x,y)dy \tag{3}$$

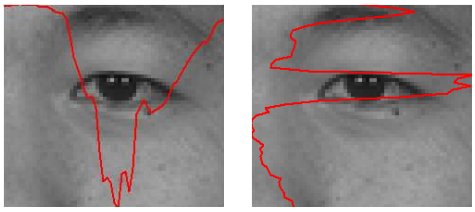
$$H_m(y)=\frac{1}{x_2-x_1}\int_{x_1}^{x_2} I(x,y)dx \tag{4}$$

$V_m(x)$, $H_m(y)$ 값은 어두운 영역에서는 값이 작게 나오며 밝은 영역에서는 값이 크게 나타나지만, 영역의 밝기 합이 같을 경우 변화를 알아내는 것이 어렵다. 이 함수를 보안하기 위해 VPF를 사용하며, VPF는 IPF를 사용하여 만들 수 있다[5]. 수직 VPF를 $\sigma_v(x)^2$, 수평 VPF를 $\sigma_h^2(y)$로 정의하면 식(5), 식(6)과 같다.

$$\sigma_v^2(x)=\frac{1}{y_2-y_1}\sum_{y_i=y_1}^{y_2}[I(x,y_i)-V_m(x)]^2 \tag{5}$$

$$\sigma_h^2(y)=\frac{1}{x_2-x_1}\sum_{x_i=x_1}^{y_2}[I(x_i,y)-H_m(y)]^2 \tag{6}$$

VPF의 특징은 해당 픽셀의 명암도가 영역의 평균 명암도보다 차이가 많이 나는 경우 크게 나타나며, 영역 픽셀의 명암도 값이 평균 명암도와 비슷하면 VPF의 값이 작게 나온다. 얼굴 영상에서 VPF 값을 분석해 보면, 평균 밝기(피부 밝기)와 차이가 많은 부분인 눈썹 부분, 눈동자 부분, 눈썹 밑 어두운 부분에서 VPF 값이 크게 나타난다.



(a) VPF_v (b) VPF_h

그림 1. 눈 영역 영상에서의 수직/수평 VPF

그림 1(a), (b)는 각각 수직, 수평 VPF를 실행한 결과이다. 영상에서 실선인 부분은 VPF 값에 해당된다. 수직 VPF의 경우 실선이 아래에 위치할수록 큰 값에 해당되며, 수평 VPF의 경우 실선이 오른쪽에 위치할수록 큰 값을 가진다. 영상에서 보이는 바와 같이 눈동자 영역에서는 평균 명암도와 차이가 크기 때문에 VPF 값이 크게 나타남을 확인 할 수 있다.

눈동자 중심좌표 (x,y)는 수직 VPF에서 가장 큰 peak 점의 x 좌표 수평 VPF에서 가장 큰 peak 점의 y 좌표가 된다. VPF만 이용하면 눈동자 뿐 아니라 눈썹 및 눈 밑의 주름부분을 검출할 가능성이 있다. 따라서 검증작업을 시행하며, 검증이 실패할 경우 원형 허프 변환을 이용하여 눈동자 중심을 찾는다.

III. 눈 영역 검증 및 2차 검출

2장에서 1차 눈동자 좌표가 검출되면, 눈동자 좌표에서 주위 8방향 이웃 픽셀의 합을 구한다. 2장에서 찾은 얼굴 영상의 가운데에서부터 위쪽으로 25×10 윈도우를 사용하여 얼굴 영역의 평균 밝기를 구한다. 25×10 윈도우에 해당하는 영역은 영상에서 코 주위의 부위에 해당된다. 이 부분의 평균 밝기를 얼굴부위의 평균 밝기로 한다. 눈동

자 좌표에서 주위 8방향 이웃 픽셀의 합이 얼굴 영역의 평균 밝기보다 작으면 눈동자를 찾은 것으로 간주한다. 그렇지 않으면 검증이 실패한 것으로 간주하고, 원형 허프 변환을 이용하여 눈동자 중심좌표를 다시 찾는다[8]. 원형 허프 변환을 이용한 수식은 식(7)과 같으며,

$$(x-a)^2+(y-b)^2=r^2 \tag{7}$$

(a,b)는 원의 중심좌표이고 r은 반지름이다.

원형 허프 변환을 사용하기 위해 왼쪽 눈, 오른쪽 눈 영역에서 다시 축소화 작업을 거친다. 검증작업이 실패한 경우는 1차 검출 시 눈썹을 검출한 경우가 많으므로 눈썹을 제거하기 위해 축소화 작업을 사용한다. 축소화 작업은 IPF의 변화 정보를 사용하여 구하며, 방법은 다음과 같다.

(1) 눈 영역 영상에서 수평 IPF값($H(y)$)과 영상의 평균 명암도(I_m)를 구한다.
(2) $H(y)$의 기울기 변화를 이용하여 I_m 값보다 작은 아래로 볼록한 구간을 구한다.
(3) 이 구간이 축소된 눈 영역 구간이 된다.

그림 2에서 보는 바와 같이 세로 흰색 선은 평균 명암도이며, 이 평균 명암도를 기준으로 아래로 볼록한 구간이 축소된 눈 영역이 된다. 구해진 축소된 얼굴 영역에서 sobel edge mask를 사용하여 에지를 추출한다. 추출된 에지는 평균 이진화를 사용하여 이진화된 이미지를 얻는다. 이진화된 이미지에서 원형 허프 변환을 이용하여 눈동자의 중심을 찾는다. 이 때, 반지름이 5픽셀인 원의 중심이 눈동자의 중심이 되며 반지름 5는 실험에 의해 검증된 값이다. 표 1은 원형 허프 변환의 r 값에 따른 결과를 나타낸다. 프로젝션 함수를 이용해서 실패한 경우, 원형 허프 변환을 이용하면 그림 3과 같이 검출할 수 있다.

표 1. 원형 허프 변환시 반경에 따른 결과 비교

반경(r)	4	5	6
검출률	89.1%	98.3%	79.3%

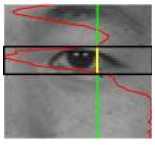
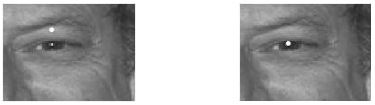


그림 2. 얼굴 영역의 축소(검은색 사각형 영역)



(a) VPF (b) 원형 허프 변환

그림 3. VPF와 원형 허프 변환을 이용한 결과

그림 3(a)의 흰색 점은 VPF를 사용하였을 때, 검출된 눈의 좌표이다. VPF를 사용하여 검출한 좌표가 검증단계에서 실패하였고, 원형 허프 변환을 사용한 결과 그림 3(b)에서와 같이 정확한 눈의 중심 좌표를 찾는다.

Ⅳ. 실험 결과 및 분석

본 논문에서는 FERET 얼굴 데이터베이스를 이용하여 얼굴 영역에서 동공의 중심점 좌표를 찾는 실험을 하였다. 사용된 얼굴 영상은 정면 영상을 사용하였으며 크기 256×384 영상을 사용하였다. 실험에 사용된 영상은 어두운 영상, 감은 눈 영상, 안경 쓴 영상이 포함되어 있다. 대부분의 검출 실패는 눈썹이 진한 경우, 눈 밑 주름이 진하게 나타나는 경우, 두꺼운 안경테를 착용하고 있는 경우, 조명의 반사가 심하게 나타나는 경우로 요약할 수 있다.

실험의 검증을 위해 실험 오차는 FERET DB의 ground truth 정보를 사용한다. ground truth 정보는 영상에서 왼쪽 눈, 오른쪽 눈에 대한 좌표를 말하며, 이 좌표는 FERET에서 제공하는 좌표이다. 이 ground truth 정보를 이용하여 수식(8)을 구한다.

$$d_{eye}=\frac{\max(\|C_l-\widetilde{C}_l\|,\|C_r-\widetilde{C}_r\|)}{\|C_l-C_r\|} \tag{8}$$

C_l, C_r은 왼쪽 눈, 오른쪽 눈 ground truth 이며, $\widetilde{C}_l$, $\widetilde{C}_r$은 제안된 알고리즘으로 검출한 왼쪽 눈, 오른쪽 눈의 좌표이다. $\|\bullet\|$ 은 Euclidean distance를 나타낸다. d_{eye} 값은 눈동자 좌표와 추출된 눈동자 좌표사이의 오차를 나타낸 값으로써, $\|C_l-C_r\|$은 눈 영역의 약 2배정도 값을 가진다. 따라서 $d_{eye}\le 0.25$의 의미는 왼쪽 눈 또는 오른쪽 눈의 오차 중 큰 값은 눈 영역 의 절반 안에 들어오는 것으로 간주할 수 있으며, $d_{eye}\le 0.05$ 오차는 입력 영상 256×384 일 때, ground truth와의 오차가 2~3 픽셀 이내에 해당된다. 본 실험에서는 $d_{eye}\le 0.25$에 해당되는 것을 성공으로 판단하였다.

표 2는 $d_{eye}\le 0.25$에서 일반영상, 감은 눈 영상, 안경 쓴 영상별로 분류한 실험결과를 나타낸다. 특히, 감은 눈에 대해서도 95.3%의 높은 검출 결과가 나타났으며, 총 검출률은 97.0%가 나왔다. 그림 4는 제안한 알고리즘으로 검출하였을 때 성공한 결과들을 나타낸 영상이다. 영상의 흰색 점이 검출한 좌표이다. 그림 5는 제안한 알고리즘으로 검출하였을 때 실패한 결과들을 나타낸 영상이다. 실패한 영상들은 눈 밑의 주름을 검출하였거나, 눈썹, 안경테를 검출하는 경우로 분류할 수 있다.

표 2. $d_{eye} \leq 0.25$ 에서의 실험결과

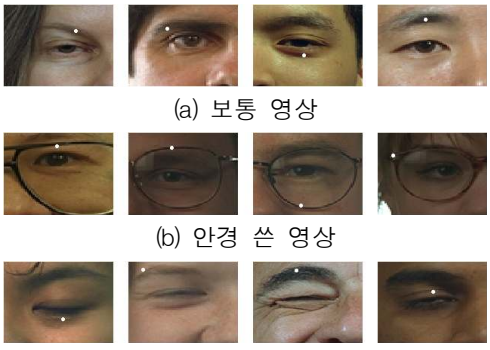
데이터베이스	검출 수/ 데이터 개수	검출률
일반영상	560 / 570	98.3%
감은 눈 영상	134 / 140	95.7%
안경 쓴 영상	92 / 100	92.0%
Total	786 / 810	97.0%

(a) 보통 영상

(b) 안경 쓴 영상

(c) 감은 눈 영상

그림 4. 다양한 영상에서의 성공한 결과



(a) 보통 영상

(b) 안경 쓴 영상

(c) 감은 눈 영상

그림 5. 다양한 영상에서의 실패한 결과

눈썹이나 눈 밑의 주름, 안경테를 검출한 경우 눈동자와 비슷한 특징이 많다. 그렇기 때문에, 1차 검출시 눈썹이나 눈 밑의 주름, 안경테가 검출되면 검증방법을 통과하게 되므로 정확한 눈동자를 검출하지 못하였다.

V. 결 론

본 논문에서는 눈 영역을 축소하는 방법과 눈 영역을 검증하는 단계를 추가하여 프로젝션 함수와 허프 변환을 결합하는 방법을 제안하였다. 먼저 프로젝션 함수를 이용하여 1차 눈동자 영역을 검출하며, 눈동자 영역의 검증작업을 거쳐 허프 변환을 이용하여 최종 눈동자 좌표를 검출하였다. 제안한 방법은 다양한 얼굴 영상에서 안정적인 성능을 보여주고 있으며, 감은 눈에서도 뛰어난 성능이 나타났고 검출한 눈동자 좌표를 검증함으로써 성능을 향상시킬 수 있었다.

눈동자 영역의 정확한 검증 알고리즘을 개발하여 눈썹, 두꺼운 안경테, 눈 밑 주름에서 검출된 영역을 수정할 수 있도록 하는 것을 추후 연구 목표로 한다.

참고문헌

[1] Y. Ming-Hsuan, D.J. Kriegman, N. Ahuja, "Detecting faces in images; a survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions,* Vol. 24, No. 1, pp.34-58, 2002.

[2] R. Gross, J. Yahng, A. Waibel, "Growing Gaussian Mixture Models for Pose Invariant Face Recognition," *Proc Int'l Conf. Pattern Recognition, V*ol. 1, pp.1088-1901, 2000.

[3] A.L. Yuille, P.W. Hallinan, D.S. Cohen, "Feature extraction from faces using deformable templates," *Int'l J. Comput. Vision 8(2),* pp. 99-111, 1992.

[4] A. Pentland, B. Moghaddam, T. Starner, "View-based and modular eigenspaces for face recognition," *Proc IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'94),* pp.84-91, 1994.

[5] G.C. Feng, P.C. Yuen, "Variance projection function and its application to eye detection for human face recognition," *Int'l J. Comput. Vision 19,* pp.899-906, 1998.

[6] P. Viola, M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," *Proc. of IEEE Conf. Computer Vision and Pattern Recognition,* pp.511-518, 2001.

[7] R. Lienhart, J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection," *Proc. IEEE ICIP,* pp.900-903, 2002.

[8] C. Kimme, D. Ballard, J. Sklansky, "Finding circles by an array of accumulators," *Comm. ACM 18, 2,* pp.120-124, 1975.

[9] T. Kawaguchi, M. Rizon, "Iris detection using intensity and edge information," *Pattern Recognition 36*, pp.549-562, 2003.